



US 20170103108A1

(19) **United States**(12) **Patent Application Publication**
DATTA et al.(10) **Pub. No.: US 2017/0103108 A1**(43) **Pub. Date: Apr. 13, 2017**(54) **EXACTLY-ONCE SEMANTICS FOR
STREAMING ANALYTICS IN
NON-IDEMPOTENT OUTPUT OPERATIONS****Publication Classification**

(51) **Int. Cl.**
G06F 17/30 (2006.01)
H04L 29/06 (2006.01)

(52) **U.S. Cl.**
CPC **G06F 17/30516** (2013.01); **H04L 65/60**
(2013.01); **G06F 17/30486** (2013.01); **G06F**
17/30339 (2013.01)

(71) Applicant: **SAMSUNG SDS AMERICA, INC.**,
Ridgefield Park, NJ (US)(72) Inventors: **Partho DATTA**, Pleasanton, CA (US);
Girish KATHALAGIRI
SOMASHEKARIAH, Santa Clara, CA
(US)(73) Assignee: **SAMSUNG SDS AMERICA, INC.**,
Ridgefield Park, NJ (US)(21) Appl. No.: **15/290,751**(22) Filed: **Oct. 11, 2016****Related U.S. Application Data**(60) Provisional application No. 62/238,812, filed on Oct.
8, 2015.(57) **ABSTRACT**

An approach for distributed stream computing in non-idempotent output operations is provided. The approach assigns an eventid to a corresponding entityid. The approach determines a minibatchid and a partitionid for a partition. The approach determines whether the partition was previously processed. The approach generates a new minibatchid and a new partitionid for a new partition based upon determining the partition was not previously processed. The approach determines whether a record was previously processed based upon determining the partition was previously processed. The approach processes the record of the partition based upon determining the record was not previously processed.

